ELSEVIER

Computational Geometry 6 (1996) 85–98

Computational Geometry
Theory and Applications

A computationally intractable problem on simplicial complexes

Ömer Eğecioğlu *, Teofilo F. Gonzalez

Department of Computer Science, University of California, Santa Barbara, CA 93106, USA

Communicated by Mikhail Atallah; submitted 18 February 1993; accepted 17 May 1995

Abstract

We analyze the problem of computing the minimum number $\text{er}(\mathcal{C})$ of internal simplexes that need to be removed from a simplicial 2-complex $\mathcal{C}$ so that the remaining complex can be nulled by deleting a sequence of external simplexes. We show that the decision version of this problem is $\mathcal{NP}$-complete even when $\mathcal{C}$ is embeddable in 3-dimensional space. Since the Betti numbers of $\mathcal{C}$ can be computed in polynomial time, this implies that there is no polynomial time computable formula for $\text{er}(\mathcal{C})$ in terms of the Betti numbers of the complex, unless $\mathcal{P} = \mathcal{NP}$. The problem can be solved in linear time for 1-complexes (graphs).

Our reduction can also be used to show that the corresponding approximation problem is at least as difficult as the one for the minimum cardinality vertex cover, and what is worse, as difficult as the minimum set cover problem. Thus simple heuristics may generate solutions that are arbitrarily far from optimal.

Keywords: Simplicial complex; Collapsing; Betti number; Algorithmic complexity; Vertex cover; Approximation algorithm

1. Introduction

We consider finite connected simplicial 2-complexes that are *pure*: i.e., all of whose maximal simplexes are 2-dimensional. Such a complex can be viewed as a collection of 2-simplexes $\mathcal{C} = \{s_1, s_2, \ldots, s_n\}$ modulo an equivalence relation that identifies pairs of simplexes s_i and s_j with $i \neq j$ along a common edge or a vertex. It is known that a simplicial 2-complex $\mathcal{C}$ has a geometric realization as a subset of the Euclidean 5-space in which each s_i is a closed triangular plane region. The reader is referred to the texts [9,13,15,17], and [18] for more information.

Recently, techniques from simplicial topology have found applications in various areas of computer science and physics. These applications include problems in distributed systems and concurrent

* Corresponding author.

SSDI 0925-7721(95)00015-1

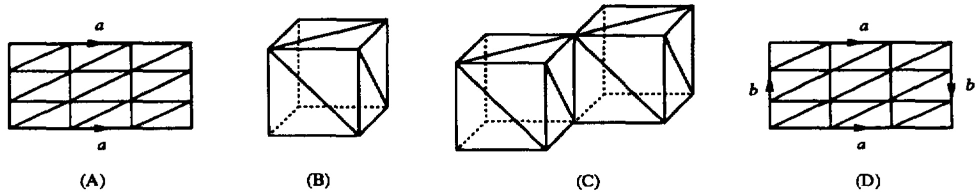


Fig. 1. (A) A cylinder segment; (B) the sphere S^2; (C) two spheres with a common edge; (D) the Klein bottle.

computation [3,12,19]; lower bound methods under the algebraic decision tree model of computation [4,5,11,20,21]; and lattice gauge theory computations in high energy physics.

In this paper we study the properties of 2-complexes in terms of the subcomplexes obtained by deleting a subset of its 2-simplexes. A 2-simplex $s \in \mathcal{C}$ is called an *external* simplex of $\mathcal{C}$ if s has at least one proper face which not shared with any other simplex in $\mathcal{C}$; otherwise s is called *internal*. Given a 2-complex $\mathcal{C}$ and a 2-simplex $s_i \in \mathcal{C}$, we denote by $\mathcal{C} - s_i$ the 2-complex obtained by restricting the given identifications defining $\mathcal{C}$ to $\{s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n\}$. We say that $\mathcal{C} - s_i$ is obtained from $\mathcal{C}$ by *removing* (*erasing*) the internal (external) simplex s_i. If $\mathcal{C}'$ is obtained from $\mathcal{C}$ by erasing an external simplex of $\mathcal{C}$, then we denote this by $\mathcal{C} \leadsto \mathcal{C}'$. More generally if two complexes $\mathcal{C}$ and $\mathcal{C}_m$ are related by a sequence of erasures of external simplexes $\mathcal{C} \leadsto \mathcal{C}_1 \leadsto \cdots \leadsto \mathcal{C}_m$, then we denote this also by $\mathcal{C} \leadsto \mathcal{C}_m$. We say that the complex $\mathcal{C}$ is *erasable* (or *nullable*) if $\mathcal{C} \leadsto \phi$. As examples, the segment of a *cylinder* in Fig. 1(A) is erasable. However the triangulation of the 2-dimensional sphere S^2 in Fig. 1(B), and the complexes in (C) and (D) are not erasable since these have no external simplexes. Note that the operation $\leadsto$ is not a topological invariant, since it can destroy the fundamental group.

Given a 2-complex $\mathcal{C}$ we define $\mathrm{er}(\mathcal{C})$ to be the minimum number of internal 2-simplexes that need to be removed from $\mathcal{C}$ so that the resulting complex is erasable. For example, for the complexes in Fig. 1(A), (B) and (C), we have $\mathrm{er}(\mathcal{C}) = 0, 1$ and 2, respectively. For these examples $\mathrm{er}(\mathcal{C})$ is equal to the second Betti number β_2 of the complex, but this is not true in general. For example the Klein bottle in Fig. 1(D) has $\beta_2 = 0$, but it has no external simplexes, and thus it does require the removal of a 2-simplex to be erasable. Furthermore it can be shown that this is not a simple anomaly due to the nonorientability of the Klein bottle.

The quantity $\mathrm{er}(\mathcal{C})$ also gives the minimum number of internal 2-simplexes that need to be removed from $\mathcal{C}$ so that the resulting complex can be *collapsed* to a lower dimensional subcomplex (see the Appendix). If $\mathcal{C}$ collapses to a d or lower dimensional subcomplex, this is denoted by $\mathcal{C} \searrow d$.

In this paper we show that the problem of computing $\mathrm{er}(\mathcal{C})$ for a given 2-complex is intractable, and furthermore, the associated approximation problem is difficult as well. More precisely, we first show that the decision version of the problem is $\mathcal{NP}$-complete. Then we establish the intractability of the corresponding approximation problem.

Erasability problem:
INSTANCE: A pair $(\mathcal{C}, k)$ where $\mathcal{C}$ is a 2-complex and k is a nonnegative integer.
QUESTION: Is $\mathrm{er}(\mathcal{C}) \leqslant k$? i.e., does $\mathcal{C}$ contain a subset $\mathcal{K}$ of 2-simplexes of cardinality at most k such that $\mathcal{C} - \mathcal{K} \leadsto \phi$?

The Erasability problem can be paraphrased as the following decision problem involving collapsibility.

Collapsibility problem:
INSTANCE: A pair $(\mathcal{C}, k)$ where $\mathcal{C}$ is a 2-complex and k is a nonnegative integer.
QUESTION: Is $\mathrm{er}(\mathcal{C}) \leqslant k$? i.e., does $\mathcal{C}$ contain a subset $\mathcal{K}$ of 2-simplexes of cardinality at most k such that $\mathcal{C} - \mathcal{K} \searrow 1$?

In the following section, we present the reduction to establish the $\mathcal{NP}$-completeness of the Erasability problem. Issues relating to the corresponding approximation problem of estimating $\mathrm{er}(\mathcal{C})$ are discussed in Section 3, and in Section 4 we conclude by considering the relationship between erasability and Betti numbers.

2. $\mathcal{NP}$-completeness of the Erasability problem

We begin by defining the *vertex cover* problem (VC). Let $G = (V, E)$ be an undirected graph with vertex set V, and edge set E. A subset $V' \subseteq V$ of vertices is said to be a vertex cover for G iff every edge in E is incident to at least one vertex in V'.

Vertex Cover problem (VC):
INSTANCE: A pair (G, k), where $G = (V, E)$ is an undirected graph, and k is an integer.
QUESTION: Does G have a vertex cover with at most k vertices?

Example 1. An instance of VC has $G = (V, E)$ depicted in Fig. 2, where $V = \{v_1, v_2, v_3, v_4\}$, $E = \{\{v_1, v_2\}, \{v_2, v_3\}, \{v_3, v_4\}, \{v_1, v_4\}, \{v_1, v_3\}\}$ and $k = 2$. This is a YES-instance of VC since $V' = \{v_1, v_3\}$ is a vertex cover with cardinality $k = 2$ for G.

VC was among the first set of problems shown to be $\mathcal{NP}$-complete [8,14]. The class of $\mathcal{NP}$-complete problems is the set of all decision problems $Q \in \mathcal{NP}$ such that SAT $\propto Q$, where SAT is the satisfiability problem defined in [8], and $\propto$ represents polynomial time reducibility [8]. The class of $\mathcal{NP}$-complete problems is very rich. We refer the reader to [6,8,14] for additional details about the

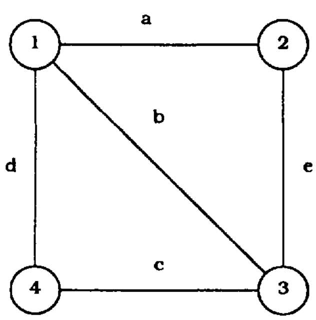


Fig. 2. Graph G of Example 1.

theory of $\mathcal{NP}$-completeness. We establish our intractability results by constructing a polynomial time reduction from VC to the Erasability problem.

2.1. Preliminaries

First we define and prove some properties of a special class of complexes that we shall use in our reduction. Fig. 3 shows a complex $\mathcal{G}$ that we refer to as the *4-gadget*. Regions that are shaded represent sets of 2-simplexes that are *not* part of the 4-gadget. The subcomplexes a_1, a_2, a_3, a_4; e_2, e_3, e_4; d, e, f and g are *cylinders*. Cylinder g is complete, except for four holes for cylinders a_1, a_2, a_3 and a_4 that attach to it, and one hole on its left end for cylinder f. Cylinders a_1, a_2, a_3 and a_4 are open on one end where they meet cylinder g. Each a_i has a hole on its side for cylinder b_i. Cylinders e_2, e_3, e_4, e and f are open on both ends. Cylinder d is complete, except for two holes on one end, and one on the other. The subcomplexes b_1, b_2, b_3 and b_4 are cylinders open on both ends. They are depicted as S-shaped regions in Fig. 3. The subcomplexes c_1, c_2, c_3 and c_4 are disks, each with two holes. The subcomplexes d_1, d_2, d_3 and d_4 are T-shaped regions with holes on their top ends, and missing bottom end as shown in Fig. 3. This complex is called the 4-gadget because there are four cylinders labeled a_1, a_2, a_3 and a_4.

In order to understand the structure of $\mathcal{G}$ better we introduce Figs 4(A), (B) and (C). Fig. 4(A) is a copy of Fig. 3 except that thin lines represent lines that are not visible from the outside. Fig. 4(B) is Fig. 4(A) after deleting the five exterior cylinders (labeled a_1, a_2, a_3, a_4 and g). Similarly, Fig. 4(C) is 4(B) after deleting the cylinders labeled f and d, and the T-shaped external regions (labeled d_1, d_2, d_3 and d_4).

We associate the special graph shown in Fig. 5 with $\mathcal{G}$. This graph contains two type of nodes: solid (s-*nodes*), and labeled (l-*nodes*). Each l-node represents a subcomplex of the 4-gadget $\mathcal{G}$ which is denoted by the same label as in Fig. 3. There are exactly three edges emanating from every s-node and each one of these edges end at an l-node. All the other edges in the graph are bidirectional and join two l-nodes. A bidirectional edge propagates erasure by the following rule.

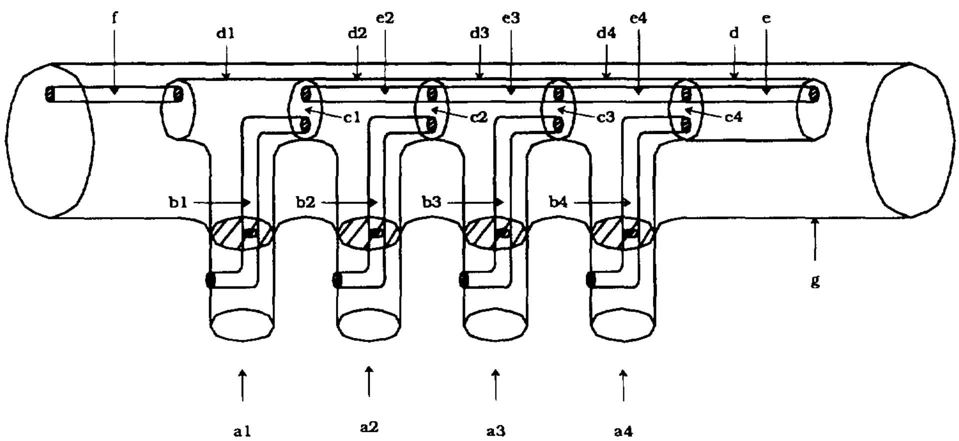


Fig. 3. 4-gadget $\mathcal{G}$.

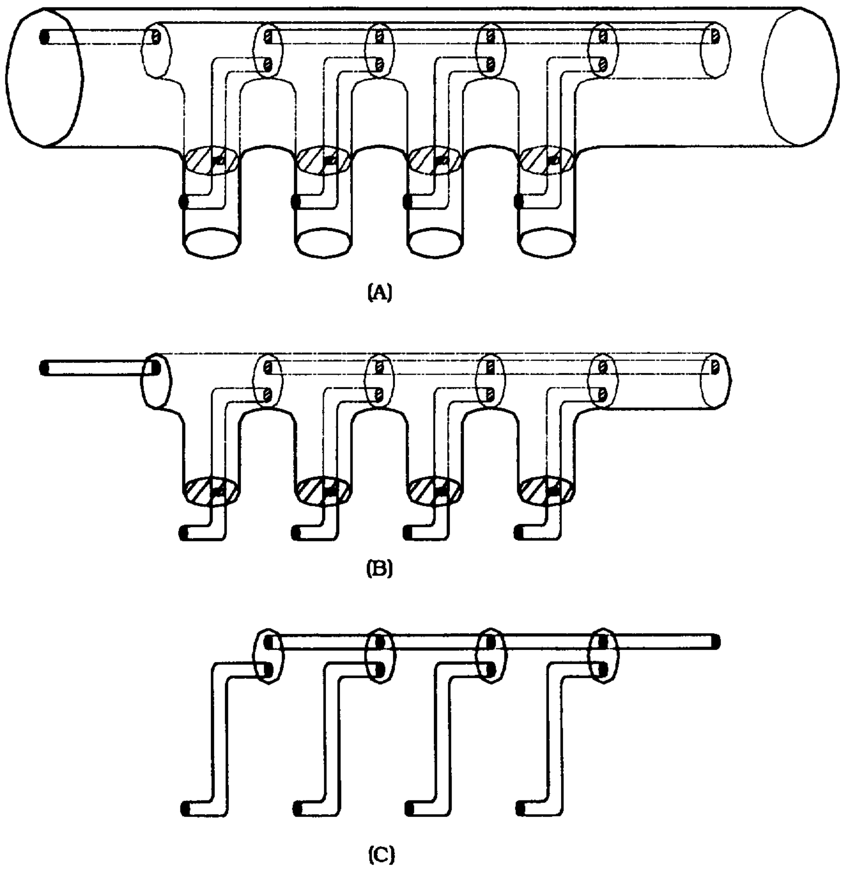


Fig. 4. Peeling of the 4-gadget $\mathcal{G}$.

Rule 1. *If one of the end nodes of a bidirectional edge is removed or erased, then the other end node can also be erased.*

For example, if the region labeled a_1 is removed or erased, then b_1 can be erased. Also, if c_1 is removed or erased, then both b_1 and e_2 can be erased. Removing or erasing c_2 will erase b_2, but not necessarily erase e_2 or e_3. Note that when a node in the graph is removed or erased, all the incoming and outgoing edges to that node are deleted. The edges emanating from s-nodes propagate erasure by the following rule:

Rule 2. *If any two l-nodes that are neighbors of an s-node are removed or erased, then the third neighbor can also be erased.*

For example, if the l-nodes d_1 and c_1 are removed or erased, then the l-node d_2 can be erased.

Given two l-nodes x, and y with a bidirectional edge between them, we use the notation $x \Rightarrow y$, to mean that node y can be erased whenever x has been removed or erased (this is an application of Rule 1). If s_x and s_x are 2-simplexes in the subcomplexes of $\mathcal{G}$ that correspond to x and y, then the notation $x \Rightarrow y$ is equivalent to $\mathcal{G} - s_x \rightsquigarrow \mathcal{G} - x - y$. Whenever there is an s-node with edges to

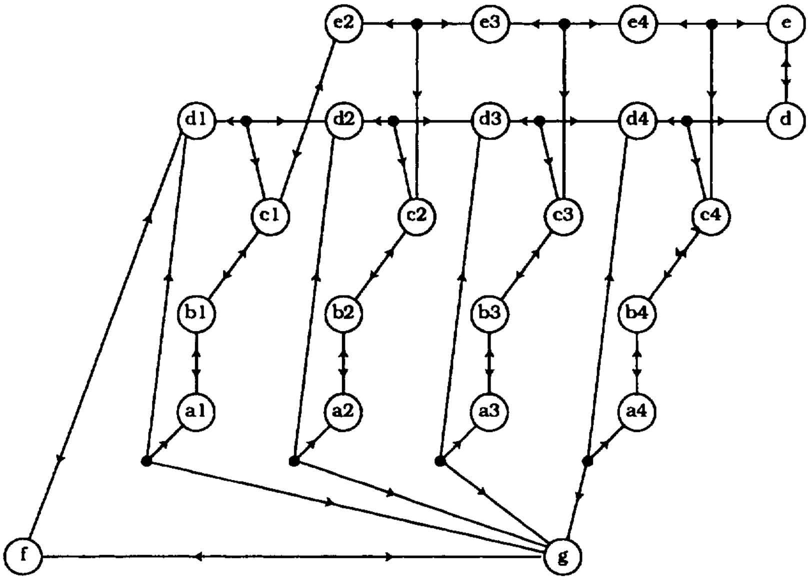


Fig. 5. Graph for the 4-gadget.

l-nodes x, y, and z, we use the notation $x \wedge y \Rightarrow z$ to mean that if x and y are erased or removed, then z can be erased. This is equivalent to $\mathcal{G} - s_x - s_y \rightsquigarrow \mathcal{G} - x - y - z$. In the graph corresponding to $\mathcal{G}$, this operation erases the l-node z and all the edges incident to it (this is an application of Rule 2).

Initially we remove a set of l-nodes. If by applying the above two rules the whole graph is erased, then the removal of a 2-simplex from each of the subcomplexes that correspond to the initially removed l-nodes in the graph suffices to erase the 4-gadget $\mathcal{G}$. Let us now observe some properties of the 4-gadget $\mathcal{G}$.

Observation 1. We note that if the l-node g is removed, then the whole graph and consequently the 4-gadget $\mathcal{G}$ is erased.

$g \Rightarrow f \Rightarrow d_1$;
$d_1 \wedge g \Rightarrow a_1 \Rightarrow b_1 \Rightarrow c_1 \Rightarrow e_2$;
$c_1 \wedge d_1 \Rightarrow d_2$;
$d_2 \wedge g \Rightarrow a_2 \Rightarrow b_2 \Rightarrow c_2$;
$c_2 \wedge e_2 \Rightarrow e_3$;
$c_2 \wedge d_2 \Rightarrow d_3$;
$d_3 \wedge g \Rightarrow a_3 \Rightarrow b_3 \Rightarrow c_3$;
$c_3 \wedge e_3 \Rightarrow e_4$;
$d_3 \wedge c_3 \Rightarrow d_4$;
$d_4 \wedge g \Rightarrow a_4 \Rightarrow b_4 \Rightarrow c_4$;
$c_4 \wedge e_4 \Rightarrow e$; and
$d_4 \wedge c_4 \Rightarrow d$.

Therefore, the removal of the l-node g erases the whole graph.

Observation 2. Another important fact that we need to establish is that if a_1, a_2, a_3 and a_4 are erased, then the whole graph (and hence the 4-gadget $\mathcal{G}$) is erased. The proof is as follows:

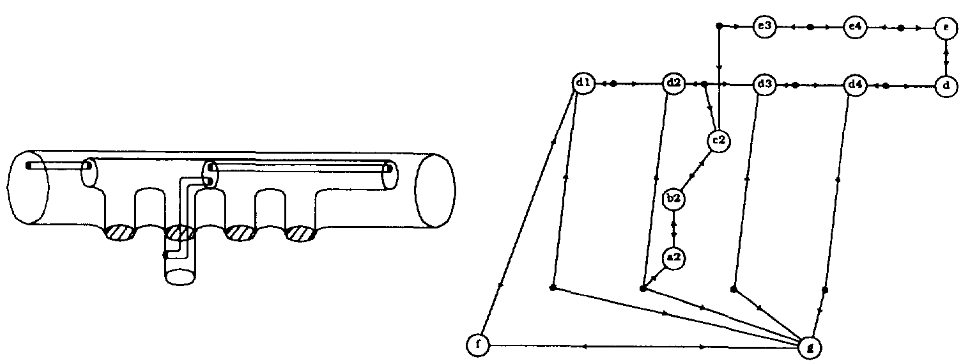
$a_1 \Rightarrow b_1 \Rightarrow c_1 \Rightarrow e_2$;
$a_2 \Rightarrow b_2 \Rightarrow c_2$;
$c_2 \wedge e_2 \Rightarrow e_3$;
$a_3 \Rightarrow b_3 \Rightarrow c_3$;
$c_3 \wedge e_3 \Rightarrow e_4$;
$a_4 \Rightarrow b_4 \Rightarrow c_4$;
$c_4 \wedge e_4 \Rightarrow e \Rightarrow d$;
$c_4 \wedge d \Rightarrow d_4$;
$c_3 \wedge d_4 \Rightarrow d_3$;
$c_2 \wedge d_3 \Rightarrow d_2$; and
$c_1 \wedge d_2 \Rightarrow d_1 \Rightarrow f \Rightarrow g$.

Therefore, if a_1, a_2, a_3 and a_4 are removed from the graph, then all of the remaining vertices can be erased.

Observation 3. Suppose that initially only three of the four l-nodes a_1, a_2, a_3 and a_4 are removed. Say these are a_1, a_2 and a_4. If s_1, s_2 and s_4 are 2-simplexes on a_1, a_2 and a_4 in $\mathcal{G}$ respectively, then $\mathcal{G} - s_1 - s_2 - s_4 \rightsquigarrow \mathcal{G}'$, where $\mathcal{G}'$ is the complex shown in Fig. 6(A). Fig. 6(B) shows the resulting graph corresponding to $\mathcal{G}'$. In this case it is simple to show that the l-nodes d_1, d_2, d_3, d_4; e_3, e_4; d, e, f and g cannot be erased by any sequence of applications of Rules 1 and 2. In fact, removal of 2-simplexes from any proper subset of the a-regions in the complex $\mathcal{G}$ is insufficient to erase $\mathcal{G}$ completely. In other words if the number of a-regions from which 2-simplexes are removed is strictly less than 4, the whole 4-gadget $\mathcal{G}$ cannot be erased.

Next we consider a generalization of the 4-gadget to the *d-gadget*. For $d \geqslant 1$, a d-gadget is the obvious extension of the 4-gadget $\mathcal{G}$ in which instead of four there are d cylinders of type a, interconnected as in Fig. 3 by using subcomplexes a_i, b_i, c_i and d_i, for $1 \leqslant i \leqslant d$.

The following lemmas concerning d-gadgets are straightforward extensions of the three observations for the case $d = 4$ given above. For brevity the proofs are omitted.



(A) (B)

Fig. 6. Resulting complex and graph after deleting a_1, a_3 and a_4.

Lemma 1. *If a 2-simplex in the g region of a d-gadget is removed, then the whole d-gadget $\mathcal{G}$ can be erased.*

Lemma 2. *If a 2-simplex from each of the a-regions of a d-gadget is removed, then the whole d-gadget $\mathcal{G}$ can be erased.*

Lemma 3. *If the number of a-regions from which 2-simplexes are removed is less than d, the whole d-gadget $\mathcal{G}$ cannot be erased.*

2.2. The reduction

Given an instance of the vertex cover problem (G, k), we construct an instance of the Erasability problem $(\mathcal{C}, k)$ by using d-gadgets as follows. Each vertex in G of degree d is represented by a d-gadget. Each cylinder $a_1, a_2, \ldots, a_d$ represents an edge incident to the vertex. Remove all the 2-simplexes at the end of all the a-regions. For each edge in G between vertices i and j, extend a cylinder connecting the end of the a-region in the d-gadget representing vertex i, to the end of an a-region in the d-gadget representing vertex j. Since the complex is constructed in 3-dimensional space, by bending cylinders appropriately all of these connections can be made. Furthermore, this reduction can be carried out in polynomial time with respect to the number of nodes and edges in G. Fig. 7 shows the instance $(\mathcal{C}, k)$ of the Erasability problem that is constructed from the graph G of Example 1 by this reduction. We claim that $\mathrm{er}(\mathcal{C}) \leqslant k$, iff G has a vertex cover of size at most k.

Lemma 4. *Suppose $\mathcal{C}$ is the 2-complex constructed from G in the above manner. Then* $\mathrm{er}(\mathcal{C}) \leqslant k$ *iff G has a vertex cover of size at most k.*

Proof. First we prove that if G has a vertex cover of size at most k, then $\mathrm{er}(\mathcal{C}) \leqslant k$. Without loss of generality, let $V' = \{v_1, v_2, \ldots, v_k\}$ be a set of k vertices in G that form a vertex cover. Now delete from $\mathcal{C}$ a 2-simplex in region g from each of the d-gadget that represent the vertices $\{v_1, v_2, \ldots, v_k\}$.

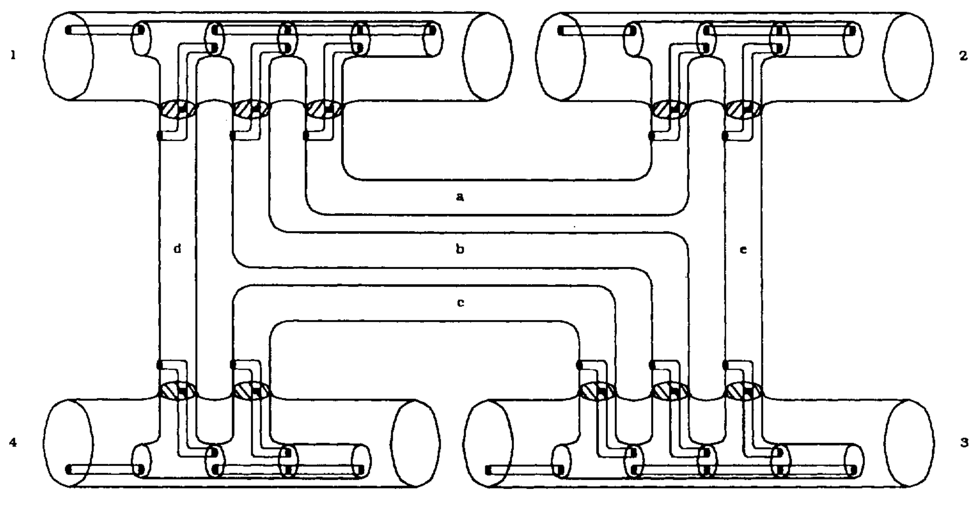


Fig. 7. Simplicial complex $\mathcal{C}$ generated for the graph G of Example 1.

By Lemma 1, each of these d-gadgets can be erased, as well as all the cylinders representing edges incident to the vertices represented by these d-gadgets. Since V' is a vertex cover for G, all the cylinders representing edges in the graph will be erased. The d-gadgets that do not represent vertices in V' can also be erased because all the edges emanating from them were erased by the above process. By Lemma 2 if all the a-regions are erased, then the d-gadget can be erased. Therefore, $\mathcal{C}$ is erased by removing a set of 2-simplexes of cardinality k, and thus $\text{er}(\mathcal{C}) \leqslant k$.

Suppose now that $\text{er}(\mathcal{C}) \leqslant k$. Assume without loss of generality that $\mathcal{K}$ is a set of 2-simplexes of cardinality k in $\mathcal{G}$ such that $\mathcal{C} - \mathcal{K} \leadsto \phi$. We can further assume that the simplexes in $\mathcal{K}$ are in the g regions of the d-gadget representing the vertices in the graph. This assumption can be made because if a 2-simplex s is removed from a cylinder representing an edge or in a region other than the g region, then by Lemma 1 one can delete a 2-simplex in the g region of the d-gadget representing one of the nodes incident to that edge or in the d-gadget where s is located, and in both cases have the net effect of erasing the 2-simplex s.

By Lemma 1 the deletion of a 2-simplex in the g region of a d-gadget results in erasing the whole d-gadget. This holds for all the d-gadgets from which we remove a 2-simplex from their g region. Let V' be the set of vertices in G that are represented by these d-gadgets. We will show that V' is a vertex cover for G. When a d-gadget is erased, then all its a-regions and their extensions are also erased. This means that the erasure propagates along the subcomplexes representing the edges adjacent to vertices in V'. A d-gadget representing a vertex in $V - V'$ can only be erased if all of its a-regions are erased (Lemmas 2 and 3). Since by assumption $\mathcal{C} - \mathcal{K} \leadsto \phi$, it must be that all the d-gadgets are erased. Consequently each edge in G must have been adjacent to a vertex in V'. Therefore, V' is a vertex cover of size k for G.

Hence, $\text{er}(\mathcal{C}) \leqslant k$ iff G has a vertex cover of size at most k. □

Since the Erasability problem is trivially in $\mathcal{NP}$, we obtain Theorem 1.

Theorem 1. *The Erasability problem is $\mathcal{NP}$-complete.*

We prove in Appendix A that the number of internal 2-simplexes that needs to be removed from $\mathcal{C}$ so that the resulting complex is collapsible to a 1-dimensional complex is also given by $\text{er}(\mathcal{C})$. Therefore,

Corollary 1. *The Collapsibility problem is $\mathcal{NP}$-complete.*

3. Approximability

Let us now consider the optimization version of the Erasability problem. By this we mean the problem of finding the least number of 2-simplexes that need to be removed in order to erase a 2-complex. Clearly, any algorithm that solves the optimization version of the Erasability problem will also solve the decision version studied in the previous section. This argument establishes that the decision version of the erasability problem is Turing reducible (see [8]) to its corresponding optimization version. Therefore, the optimization version of the Erasability problem is an $\mathcal{NP}$-hard problem. Since the optimization version is computationally intractable, we turn our attention to the problem of generating suboptimal solutions, i.e., study the corresponding approximation problem.

An approximation algorithm has an approximation bound of c for a given problem Q, if for every problem instance I of Q the algorithm generates solutions with objective function value $\hat{f}_I \leqslant cf_I^*$, where f_I^* is the objective function value of an optimal solution to problem instance I. Our reduction from VC to the Erasability problem given in the previous section is *approximation preserving*. This is because the simplicial complex generated by our reduction can be erased by removing a subset $\mathcal{K}$ of 2-simplexes of cardinality at most k iff the corresponding graph has a vertex cover V' with cardinality at most k. Furthermore, the vertex cover V' can be identified quickly from $\mathcal{K}$. An interesting consequence of this property is that any approximation algorithm for the optimization version of the Erasability problem (henceforth we refer to this problem as the Erasability approximation problem) with approximation bound c, is also an approximation algorithm for VC with an identical approximation bound and time complexity bound (modulo the time complexity of the reduction, which is minimal).

An approximation algorithm with approximation bound $1+\varepsilon$, for every $\varepsilon > 0$, that takes polynomial time with respect to the input size n, and $1/\varepsilon$ is called a *fully polynomial time approximation scheme*. If the algorithm takes time polynomial time with respect to the parameter n only, then it is called a *polynomial time approximation scheme*. Since the VC is a strongly $\mathcal{NP}$-complete problem (see [8]), it cannot have a fully polynomial time approximation algorithm unless $\mathcal{P} = \mathcal{NP}$. Because of our approximation preserving reduction from VC to the Erasability problem, it follows that unless $\mathcal{P} = \mathcal{NP}$, there can be no fully polynomial time approximation algorithm for the Erasability problem.

In [10] several approximation algorithms for VC with an approximation bound of 2 are discussed. Unfortunately, none of these algorithms can be generalized to the Erasability approximation problem. In the worst case, simple heuristic rules generate arbitrarily bad solution to VC, therefore they also generate arbitrarily bad solutions to the Erasability problem. Other heuristics for the Erasability problem that we analyzed also generate arbitrarily bad solutions in the worst case. The above observations imply that the Erasability approximation problem seems harder than the one for VC. We now present a rigorous argument to support this statement. We begin by defining the Set Cover problem which is a generalization of VC.

Set Cover problem (SC):
INSTANCE: Finite collection of subsets $S = (S_1, S_2, \ldots, S_m)$ of a finite set U, and an integer k.
QUESTION: Are there k sets in S whose union is U?

SC is an $\mathcal{NP}$-complete problem, since it is an obvious generalization of VC. The optimization version of SC asks for the least number of the sets in S whose union is U. The main reason we failed in trying to find a polynomial time approximation algorithm with a constant approximation bound for the Erasability approximation problem is that SC also polynomially reduces to the Erasability problem, and the reduction is approximation preserving. Our reduction is as follows: Represent each set with d elements by a d-gadget, and remove all the s-simplexes at the end of the a-regions. Each a-region is used to represent one of the elements in the set corresponding to the gadget. All a-cylinders representing the same object are extended and joined together in 3-dimensional space by bending them appropriately. It is simple to show that this reduction takes polynomial time. One can also establish that the reduction is approximation preserving, because the simplicial complex generated by this reduction can be erased by removing a subset $\mathcal{K}$ of 2-simplexes of cardinality at most k iff the corresponding instance of SC has a set cover of cardinality at most k, and such a cover can be identified quickly from $\mathcal{K}$. Therefore, the Erasability approximation problem is as hard as the approximation problem for SC. We have not been able to construct an approximation preserving reduction from the SC to the Erasability problem. That is why the Erasability approximation problem seems harder than the one for the set cover.

The non-approximability result in [1] establishes that the SC does not have a polynomial time approximation scheme unless $\mathcal{P} = \mathcal{NP}$. Because of our approximation preserving reduction, this result translates directly to the Erasability problem. More recently a stronger result was reported in [2] for the SC problem where it is shown that SC cannot be approximated in polynomial time within *any* constant ratio unless $\mathcal{P} = \mathcal{NP}$. By our approximation preserving reduction, it follows that the Erasability approximation problem cannot be solved in polynomial time within any constant ratio unless $\mathcal{P} = \mathcal{NP}$. Other non-approximability results for the SC problem [2] and [16] also translate directly to the Erasability problem.

4. Relationship of $\mathrm{er}(\mathcal{C})$ with Betti numbers

The Betti numbers β_i for $0 \leqslant i \leqslant d$ of a d-dimensional simplicial complex $\mathcal{C}$ are topological invariants related to high dimensional connectivity properties of $\mathcal{C}$. The Betti number β_0 is the number of connected components of $\mathcal{C}$, and intuitively, β_i is the number of "i-dimensional holes" in $\mathcal{C}$ (see [13]). Since $\mathrm{er}(\mathcal{C})$ appears to count the number of three dimensional regions enclosed by $\mathcal{C}$, at first it seems reasonable to expect some relationship between $\mathrm{er}(\mathcal{C})$ and the Betti numbers of $\mathcal{C}$. To this end, we first consider 1-dimensional simplicial complexes, also referred to as *graphs*. For a connected graph $\mathcal{G}$ with n vertices and e edges (1-simplexes), the 1-dimensional Betti number β_1 is the maximum number of linearly independent elementary cycles in $\mathcal{G}$ (see [15]). Equivalently, β_1 is the dimension of the circuit space of $\mathcal{G}$. For graphs, β_1 and the rank r of the incidence matrix of $\mathcal{G}$ are related by $\beta_1 = e - r$, and thus β_1 can be found by a rank computation in polynomial time. Actually the 1-dimensional Betti number for graphs can be expressed explicitly by $\beta_1 = e - n + 1$ as a consequence of the the Euler–Poincaré relation: for an arbitrary d-dimensional complex, this relation is

$$\chi(\mathcal{C}) = \sum_{i=0}^{n} (-1)^i \alpha_i = \sum_{i=0}^{n} (-1)^i \beta_i,$$

where α_i be the number of i-simplexes of $\mathcal{C}$. This common value is the Euler characteristic of $\mathcal{C}$.

It is known that all of the Betti numbers $\beta_0, \beta_1, \ldots, \beta_d$ of a general d-dimensional complex $\mathcal{C}$ can be computed from the quantities α_i, and the ranks of the incidence matrices relating the i-dimensional simplexes of $\mathcal{C}$ to its $(i-1)$-dimensional simplexes, $1 \leqslant i \leqslant n$. Consequently these invariants can be computed in polynomial time in the total number of simplexes of $\mathcal{C}$ (see [13,15,18]).

Note that when we restrict the operation of removal and erasure to 1-dimensional simplexes, then the notions of Erasability and collapsibility coincide. If $\mathcal{G}$ is a graph then $\mathrm{er}(\mathcal{G}) = \beta_1$. Therefore the Erasability problem for graphs is in $\mathcal{P}$. The quantity $\mathrm{er}(\mathcal{C})$ and the Erasability problem can be defined for higher dimensional complexes by extending the notions of internal and external simplexes of $\mathcal{C}$ in the obvious fashion. However the intractability of the Erasability problem for 2-complexes implies that the general problem for arbitrary $d \geqslant 2$ dimensions is necessarily $\mathcal{NP}$-complete. Furthermore, since the Betti numbers can be computed in polynomial time, this also implies that there can be no computationally easy formula that relates $\mathrm{er}(\mathcal{C})$ to the Betti numbers of the complex. More precisely, unless $\mathcal{P} = \mathcal{NP}$, there can be no polynomial time computable function $f(x_0, x_1, \ldots, x_d)$ for which $\mathrm{er}(\mathcal{C}) = f(\beta_0, \beta_1, \ldots, \beta_d)$.

We note that the Erasability problem is solvable in polynomial time for constant k, since it suffices to generate all k-element subsets $\mathcal{K}$ of internal simplexes of $\mathcal{C}$ and for each subset $\mathcal{K}$ check in polynomial time whether $\mathcal{C} - \mathcal{K} \rightsquigarrow \phi$.

5. Remarks

A preliminary version of the $\mathcal{NP}$-completeness result presented in this paper, in which a restricted version of SAT is reduced directly to the Erasability problem using nonorientable components appears in [7].

Appendix A

Suppose s is an external simplex of a 2-complex $\mathcal{C}$. Assume $s = abc$ with vertices (a, b, c) and 1-simplexes (ab, ac, bc). Assume that the face bc is not shared by any other 2-simplex of $\mathcal{C}$. The operation of going from $\mathcal{C}$ to $\mathcal{C}' = \mathcal{C} - s + ab + ac$ (where $+$ is union) is called an *elementary collapsing*, denoted by $\mathcal{C} \searrow \mathcal{C}'$ (see [9, p. 49]). Geometrically, the complex $\mathcal{C}'$ is a deformation retract of $\mathcal{C}$ obtained by "pushing in" on the free face bc of s. If ab or ac becomes an external 1-simplex in $\mathcal{C}'$ after collapsing bc (i.e., vertex b or vertex c is not shared with any other 1-simplex in $\mathcal{C}'$) then we can collapse these as well, so that no external 1-simplexes are left in the resulting complex. Here we will use the term elementary collapsing as this composite operation of collapsing (i.e., a 2-complex followed by at most two 1-complex collapsings), and use the symbol $\mathcal{C} \searrow \mathcal{C}'$ for it. If there is a finite sequence of elementary collapsings $\mathcal{C} \searrow \mathcal{C}_1 \searrow \cdots \searrow \mathcal{C}_k$ then we say that $\mathcal{C}$ *collapses* to $\mathcal{C}_k$, written $\mathcal{C} \searrow \mathcal{C}_k$. If $\mathcal{C}$ collapses to a subcomplex of dimension d or less, then this is denoted by $\mathcal{C} \searrow d$. A complex $\mathcal{C}$ with $\mathcal{C} \searrow 0$ is called *collapsible* [9]. As examples, the complex in Fig. 1(A) collapses to the circle S^1 and thus $\mathcal{C} \searrow 1$. The complexes in Fig. 1(B) and (C) do not collapse to any of their subcomplexes.

Proposition 1. *Let* $\mathrm{co}(\mathcal{C})$ *denote the minimum number of* 2*-simplexes* $\mathcal{K}$ *that need to be removed from* $\mathcal{C}$ *so that* $\mathcal{C} - \mathcal{K} \searrow 1$. *Then* $\mathrm{co}(\mathcal{C}) = \mathrm{er}(\mathcal{C})$.

Proof. First we compare the effects of the operations of collapsing and erasing of an external 2-simplex. In Fig. 8, the dark subsimplexes marked in column one are assumed to be shared with other 1- or 2-dimensional simplexes in $\mathcal{C}$. The result of erasing and collapsing the external 2-simplex in column one are given in columns two and three, respectively.

Suppose $\mathcal{C}$ has n 2-simplexes, and assume $k = \mathrm{er}(\mathcal{C})$. Suppose $\mathcal{K}$ is a cardinality k set of 2-simplexes of $\mathcal{C}$ such that $\mathcal{C} - \mathcal{K} \rightsquigarrow \phi$. Let $\mathcal{C}_1 = \mathcal{C} - \mathcal{K}$. Since $\mathcal{C}_1 \rightsquigarrow \phi$, the 2-simplexes of $\mathcal{C}_1$ can be ordered as $e_1, e_2, \ldots, e_m$ with $k + m = n$, such that e_1 is an external simplex of $\mathcal{C}_1$ and if $\mathcal{C}_{i+1} = \mathcal{C}_i - e_i$ then e_{i+1} is an external simplex of $\mathcal{C}_{i+1}$ for $i = 1, 2, \ldots, m-1$. In other words, we have a sequence

$$\mathcal{C}_1 \rightsquigarrow \mathcal{C}_2 \rightsquigarrow \cdots \rightsquigarrow \mathcal{C}_m \rightsquigarrow \mathcal{C}_{m+1}, \tag{1}$$

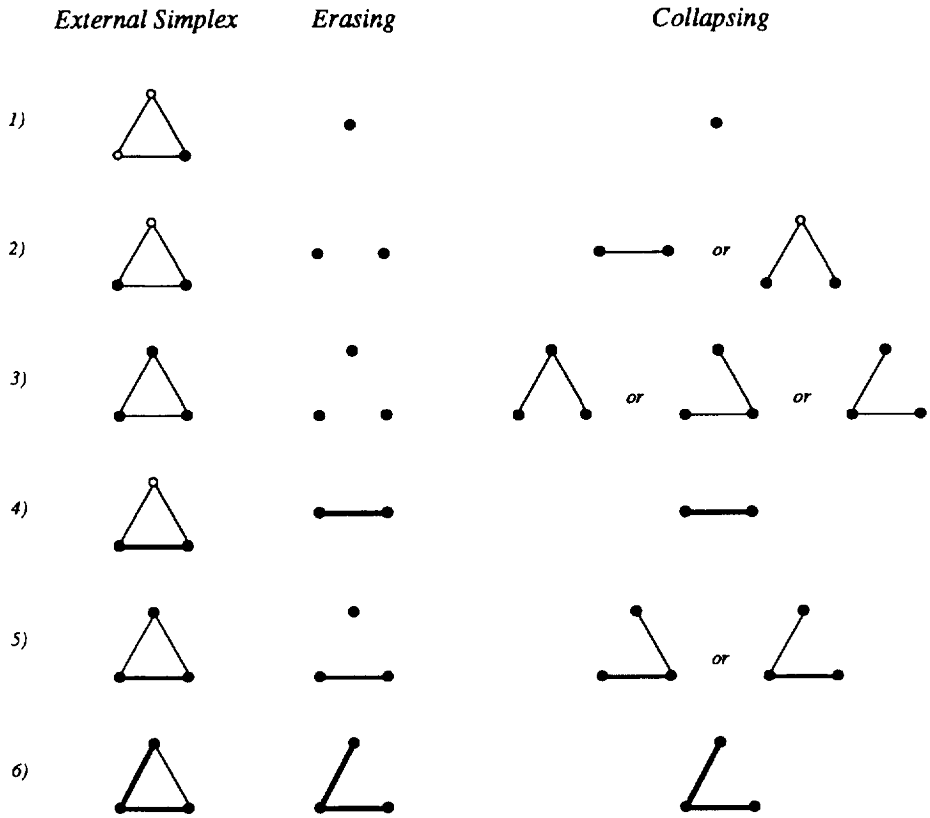


Fig. 8. Erasing versus collapsing external simplexes.

where $\mathcal{C}_{m+1} = \phi$ and the external 2-simplex e_i of $\mathcal{C}_i$ is erased in going from $\mathcal{C}_i$ to $\mathcal{C}_{i+1}$. Let $\mathcal{D}_1 = \mathcal{C}_1$. We claim that $\mathcal{D}_1 \searrow 1$ by a sequence of the form

$$\mathcal{D}_1 \searrow \mathcal{D}_2 \searrow \cdots \searrow \mathcal{D}_m \searrow \mathcal{D}_{m+1}, \tag{2}$$

in which e_i is an external simplex of $\mathcal{D}_i$, and we go from $\mathcal{D}_i$ to $\mathcal{D}_{i+1}$ by collapsing e_i, $i = 1, 2, \ldots, m$. Since each collapsing in such a sequence eliminates a 2-simplex from $\mathcal{C}$, the final complex $\mathcal{D}_{m+1}$ contains no 2-simplexes. Therefore we only need to show that e_i is an external simplex of $\mathcal{D}_i$, $i = 1, 2, \ldots, m$. However by induction on m, this fact is a consequence of the comparison between collapsing and erasing given in Fig. 8. Therefore $\mathrm{er}(\mathcal{C}) = k \leqslant \mathrm{co}(\mathcal{C})$. The steps of going from the $\mathcal{C}_i$'s to $\mathcal{D}_i$ by a sequence of the form (2) can also be reversed. A set of 2-simplexes $\mathcal{K}$ of cardinality $k = \mathrm{co}(\mathcal{C})$ of $\mathcal{C}$ such that $\mathcal{C} - \mathcal{K} \searrow 1$ gives rise to a sequence of erasures of the form (1) in which the 2-simplexes that are collapsed in each step of (2) are now erased instead in (1). From Fig. 8, going from collapsing to erasing deletes all of the 1-simplexes left under collapsing. Therefore $\mathcal{C}_{m+1} = \phi$ and $\mathcal{C} - \mathcal{K} \rightsquigarrow \phi$. Thus $\mathrm{er}(\mathcal{C}) \geqslant \mathrm{co}(\mathcal{C})$. $\square$

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan and M. Szegedy, Proof verification and intractability of approximation problems, Proc. 31st Symposium on Foundations of Computer Science (1992).

[2] M. Bellare, S. Goldwasser, C. Lund and A. Russell, Efficient probabilistically checkable proofs and applications to approximation, Proc. 25th Annual Symposium on the Theory of Computing (1993) 294–304.

[3] E. Borowsky and E. Gafni, Generalized FLP impossibility result for t-resilient asynchronous computations, Proc. 25th Annual Symposium on the Theory of Computing (1993) 91–100.

[4] A. Björner and L. Lovász, Linear decision trees, subspace arrangements and Möbius functions, Preprint, 1992.

[5] A. Björner, L. Lovász and A. C. Yao, Linear decision trees: volume estimates and topological bounds, Proc. 24th Annual Symposium on the Theory of Computing (1992) 170–177.

[6] S.A. Cook, The complexity of theorem proving procedures, Proc. 3rd Annual Symposium on the Theory of Computing, Association for Computing Machinery, New York (1971) 151–158.

[7] Ö. Eğecioğlu and T.F. Gonzalez, A computationally intractable problem on simplicial complexes, Tech. Rep. UCSB TRCS93-1, January 1993.

[8] M.J. Garey and D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness (W.H. Freeman and Company, New York, 1979).

[9] L.C. Glaser, Geometrical Combinatorial Topology, Vol. I (Van Nostrand Reinhold Company, New York, 1970).

[10] T.F. Gonzalez, LP-free approximation algorithm for the minimum weight vertex cover problem, Information Processing Letters 54 (1995) 129–131.

[11] D. Grigoriev, M. Karpinski and N. Vorobjov, Complexity lower bounds on testing membership to a polyhedron by algebraic decision trees, Proc. 26th Annual Symposium on the Theory of Computing (1994) 635–643.

[12] M. Herlihy and N. Shavit, The asynchronous computability theorem for t-resilient tasks, Proc. 25th Annual Symposium on the Theory of Computing (1993) 111–120.

[13] J.G. Hocking and G.S. Young, Topology (Addison-Wesley, Reading, MA, 1961).

[14] R.M. Karp, Reducibility among combinatorial problems, in: R.E. Miller and J.W. Thatcher, eds., Complexity of Computer Computations (Plenum, 1972) 85–103.

[15] S. Lefschetz, Introduction to Topology (Princeton University Press, New Jersey, 1949).

[16] C. Lund and M. Yannakakis, On the hardness of approximating minimization problems, Proc. 25th Annual Symposium on the Theory of Computing (1993) 286–293.

[17] W.S. Massey, Algebraic Topology: An Introduction (Harcourt-Brace, New York, 1967).

[18] J.R. Munkres, Elements of Algebraic Topology (Addison-Wesley, Menlo Park, 1984).

[19] M. Saks and F. Zaharoglou, Wait-free k-set agreement is impossible: the topology of public knowledge, Proc. 25th Annual Symposium on the Theory of Computing (1993) 101–110.

[20] A.C. Yao, Algebraic decision trees and Euler characteristics, Proc. 33rd Symposium on Foundations of Computer Science (1992) 268–277.

[21] A.C. Yao, Decision tree complexity and Betti numbers, Proc. 26th Annual Symposium on the Theory of Computing (1994) 615–624.